Increased hazard of mortality in cases compatible with SARS-CoV-2 variant of concern 202012/1 - a matched cohort study

Robert Challen [1,2,3]; Ellen Brooks-Pollock[3,4,5,]; Jonathan M Read[3,6]; Louise Dyson[3,7]; Krasimira Tsaneva-Atanasova[1,8]; Leon Danon [3,5,8,9]

1) College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, Devon, UK.
2) Somerset NHS Foundation Trust, Taunton, Somerset, UK.
3) Joint Universities Pandemic and Epidemiological Research (JUNIPER) consortium.
4) University of Bristol, Bristol Veterinary School, Langford, Bristol, UK.
5) Bristol Medical School, Population Health Sciences, University of Bristol, Bristol, UK.
6) Lancaster Medical School, Lancaster University, Bailrigg, Lancaster, UK.
7) The Zeeman Institute for Systems Biology & Infectious Disease Epidemiology Research, School of Life Sciences and Mathematics Institute, University of Warwick, Coventry, UK.
8) The Alan Turing Institute, British Library, 96 Euston Rd, London, UK.
9) Department of Engineering Mathematics, University of Bristol, UK.

Abstract:

Objectives - To establish whether there is any change in mortality associated with infection of a new variant of SARS-CoV-2 (VOC-202012/1), first detected in UK in December 2020, compared to that associated with infection with circulating SARS-CoV-2 variants.

Design - Matched cohort study. Cases are matched by age, gender, ethnicity, index of multiple deprivation, lower tier local authority region, and sample date of positive specimen, and differing only by detectability of the spike protein gene using the TaqPath assay - a proxy measure of VOC-202012/1 infection.

Setting - United Kingdom, Pillar 2 COVID-19 testing centres using the taqPath assay.

Participants - 54,773 pairs of participants testing positive for SARS-CoV-2 in Pillar 2 between 1st October 2020 and 29th January 2021.

Main outcome measures - Death within 28 days of first positive SARS-CoV-2 test.

Results - There is a high probability that the risk of mortality is increased by infection with VOC-202012/01 ($p < 0.001$). The mortality hazard ratio associated with infection with VOC-202012/1 compared to infection with previously circulating variants is 1.7 (95% CI 1.3 - 2.2) in patients who have tested positive for COVID-19 in the community. In this comparatively low risk group, this represents an increase of deaths from 1.8 in 1000 to 3.1 in 1000 detected cases.

Conclusions - If this finding is generalisable to other populations, VOC-202012/1 infections have the potential to cause substantial additional mortality over and previously circulating variants. Healthcare capacity planning, national and international control policies are all impacted by this finding, with increased mortality lending weight to the argument that further coordinated and stringent measures are justified to reduce deaths from SARS-CoV-2.

Keywords : COVID-19; SARS-CoV-2; B.1.1.7; VOC-202012/01; mortality

Competing interests: No financial relationships with any organisations that might have an interest in the submitted work in the previous three years, no other relationships or activities that could appear to have influenced the submitted work.

Funding: RC and KTA gratefully acknowledge the financial support of the EPSRC via grants EP/N014391/1, EP/T017856/1, and NHS England, Global Digital Exemplar programme. LDanon and KTA gratefully acknowledge the financial support of The Alan Turing Institute under the EPSRC grant EP/N510129/1. LDanon, RC and EBP are supported by MRC (MC/PC/19067) . JMR acknowledges support from EPSRC (EP/N014499/1) and MRC (MR/S004793/1, MR/V028456/1). EBP was partly supported by the NIHR Health Protection Research Unit in Behavioural Science and Evaluation at University of Bristol, in partnership with Public Health England (PHE). LDanon, EBP, JMR and LDyson are further supported by MRC (MR/V038613/1) and LDanon by EPSRC EP/V051555/1. LDyson also received support through the MRC through the COVID-19 Rapid Response Rolling Call (grant number MR/V009761/1).

Authors contributions: RC and LDanon designed the study, all authors discussed the concept of the article and RC performed analyses and wrote the initial draft. KTA, EB-P, LDanon, JMR and LDyson commented and made revisions. All authors read and approved the final manuscript. RC is the guarantor. The views presented here are those of the authors and should not be attributed to Somerset Foundation NHS Trust or the GDE.

Acknowledgements: We extend our thanks to Professor David Spiegelhalter for comments on an early draft, and to Dr Nick Gent at Public Health England for providing access to the data.

Patient and public involvement: Due to the nature of this research there was no involvement of patients or members of the public in the design or reporting of this study. Direct dissemination to study participants is not possible.

Data sharing: Data are available to members of universities that have data sharing agreements in place with Public Health England.

Ethics: The data were supplied from the SGSS database and death reports after anonymisation under strict data protection protocols agreed between the University of Exeter and Public Health England. The ethics of the use of these data for these purposes was agreed by Public Health England with the Governments SPI-M(O) / SAGE committees. The research was assessed as not needing NHS Research Ethics Committee review.

Increased hazard of mortality in cases compatible with SARS-CoV-2 variant of concern 202012/1 - a matched cohort study

Introduction

A variant of concern of the SARS-CoV-2 virus [1] (VOC-202012/1, variant B.1.1.7 - 'new variant') was identified in December 2020 in the South East of the United Kingdom. It spread rapidly from there to London and the rest of the UK, with three quarters of infections being attributable to the new variant by 31st December 2020 [2], suggesting that the variant spreads more easily. The UK implemented a second national lockdown (5/11/2020 - 2/12/2020) which coincided with the relative growth of VOC-202012/1. Following the lockdown additional control measures were implemented as the increased rate of spread of the new variant became apparent and was made public [3]. International restrictions on travel from the UK quickly followed, in particular to France and the rest of Europe late in December 2020 to curb spread of the new variant to other countries, despite evidence that it was already present outside the UK. Since then, VOC-202012/1 has been observed to be increasing in prevalence in Europe [4], and the US [5].

Conveniently, multiplex target Polymerase Chain Reaction (PCR) tests used in the parts of the UK national testing system are able to distinguish VOC-202012/1 from other variants. When tested using the Thermo TaqPath system it has been shown that in the UK there is a close correlation between VOC-202012/1 cases confirmed by sequencing and the S-gene target failure (SGTF) of detection, over the time period of this study. SGTF cases have subsequently been used as a proxy to track the progression of this variant [8–10,2].

Sequencing VOC-202012/1 revealed 14 genetic mutations, 8 of which occurred in parts of the genome that codes for the spike protein, responsible for cell binding [6]. These mutations appear to have imparted a phenotypic change to the cell binding mechanism [7–10,2], with the potential for increased infectivity [11,12]. The impact of the change on clinical presentation, patient outcome and, most importantly mortality, remains poorly understood.

Here, using linked data from syndromic community testing and death records, we assess whether the new SARS-CoV-2 variant is associated with an different hazard of mortality compared to the older variants.

Methods

The study primarily set out to determine if there is a change in mortality in patients testing positive for SARS-CoV-2 and with PCR test results compatible with VOC-202012/1. This question is difficult to answer as during the period under study, the rates of COVID-19 cases in the UK have increased dramatically, putting hospital services under strain, which in turn affects mortality [13] and potentially biases observations of mortality.

Study design

We conducted a matched cohort study, using an incidence density design. COVID-19 incidence and burden on hospitals varies in space and time and to address this as a potential source of bias, we match closely to time and location of cases, and assess the variability of our estimates when relaxing the matching criteria.

Inclusion criteria

Individuals who had a single positive COVID-19 test result, during the period from 1st October 2020 until 29th January 2021 were included. These were restricted to test results that reported a PCR cycle threshold (CT) value. This subset of tests were community samples (Pillar 2) processed in the high-throughput "Lighthouse" laboratories that employ the Thermo TaqPath COVID-19 multiplex Polymerase Chain Reaction (PCR) assay which amplifies the open reading frame 1a/b junction (ORF1ab) and the N and S-genes of the SARS-CoV-2. We included people that had a single positive PCR test using the TaqPath assay and for which PCR cycle threshold (CT) values for the S, N and ORF1ab components of SARS-CoV-2 were available. Data were collected by Public Health England into a centralised database detailing the type of test performed and the results.

Data processing

We classified the results as S+N+ORF1ab+ ("S-gene positive" - compatible with previous variants) for results that had the following CT values: S-gene < 30; N gene < 30; ORF1ab gene < 30. We classified S-N+ORF1ab+ ("S-gene negative" - compatible with VOC-202012/1) for results that had CT values: S-gene not detected; N-gene < 30; ORF1ab gene < 30. All other combinations of known CT values were classified as "Equivocal" and excluded from further analysis.

We linked the line list of case details and line list of details of death (if the patient died) using a unique study identifier. Cases uninformative for S-gene status, were classified as "Unknown", and were also excluded; these are generally samples not processed in Lighthouse laboratories, and include hospital cases.

During the study period hospitals experienced a period of intense demand in areas where there were large outbreaks of VOC-202012/01, which potentially adversely impacted patient outcomes. To control for any systematic bias this could have introduced, we paired individuals with S-gene positive test results (controls) to individuals with S-gene negative test results (cases - highly likely to be VOC-202012/01) by matching on gender, ethnicity, index of multiple deprivation (IMD), location (as lower tier local authority region ~ 190,000 people), age (within a tolerance of ±5 years), and date of positive specimen (within a tolerance of ±1 day). Before pairing we excluded all patients less than 30 years of age as they did not contribute to the mortality data. Some cases matched multiple controls and vice versa, so we sampled the cases and controls randomly within our framework to generate 50 replicates, ensuring no case or control was present more than once in each replicate.

We compared the rates of death within 28 days of a positive COVID-19 test in Pillar 2 data between S-gene positive controls versus S-gene negative cases. We calculated the hazard ratio of

death given a S-gene negative test result, versus death given a S-gene positive test result using a Cox proportional hazards model [14] with age (years) as a linear covariate, taking into account censoring, and directly using partial maximum likelihood estimators from paired data as described by Shinozaki et al.[15]. All analyses were performed in R (version 3.6.3)[16–18] and all analysis code is available at doi: 10.5281/zenodo.4501299.

Sensitivity analyses

We examined different inclusion criteria for sources of systematic bias. We systematically adjusted values for cycle thresholds for the S, N and ORF1ab genes, and the tolerances of our algorithm to match both inexact age and inexact specimen dates.

Results

At the time of analysis there were 3.3M patients in the UK who have at some point tested positive for SARS-CoV-2 by PCR. Out of these there were 932,047 patients older than 30 with positive TaqPath tests between 1st Oct 2020 and 28th January 2021. From these we identified 214,365 matching pairs of patients with similar age and specimen date, and identical gender, ethnicity, geography, index of multiple deprivation, and differing only by S-gene status. Sampling these pairs to ensure they represented unique patients resulted in 50 replicates with an average of 54,773 cases and 54,773 controls per replicate. Of these average 109,545 patients, 272 died within 28 days of a positive test (0.24%) - see Table 1. The matching and sampling process is observed to control well for all demographic variables considered, and geographic variables (with slight mismatches due to differences in scale from matching and reporting). With age and specimen date, where we allowed small tolerances, the average difference between ages in the S-gene positive and S-gene negative arms was 0.0 years and a mean difference of 0.2 days for specimen date (with S negative specimens taken later than S positives).

Compared to cases we observe that people who go on to die are generally older (mean 65.4 years old in deaths versus 46.2 years old in all cases), and a higher proportion are men, as has been reported by previous analysis [19]. We note both cases and deaths are under-represented in the South West and East of England where the Pillar 2 labs have not used TaqPath assays until recently.

We found an average of 171 deaths out of 54773 patients in the S-gene negative arm of the study compared to 101 out of 54773 in the S-gene positive control arm. This gives a hazard ratio of 1.7 (95% confidence intervals 1.3 - 2.2; $p = <0.001$) - see table 2. The rate of death of S-gene negative and S-gene positive cases over time diverges after 14 days, shown in figure 2.

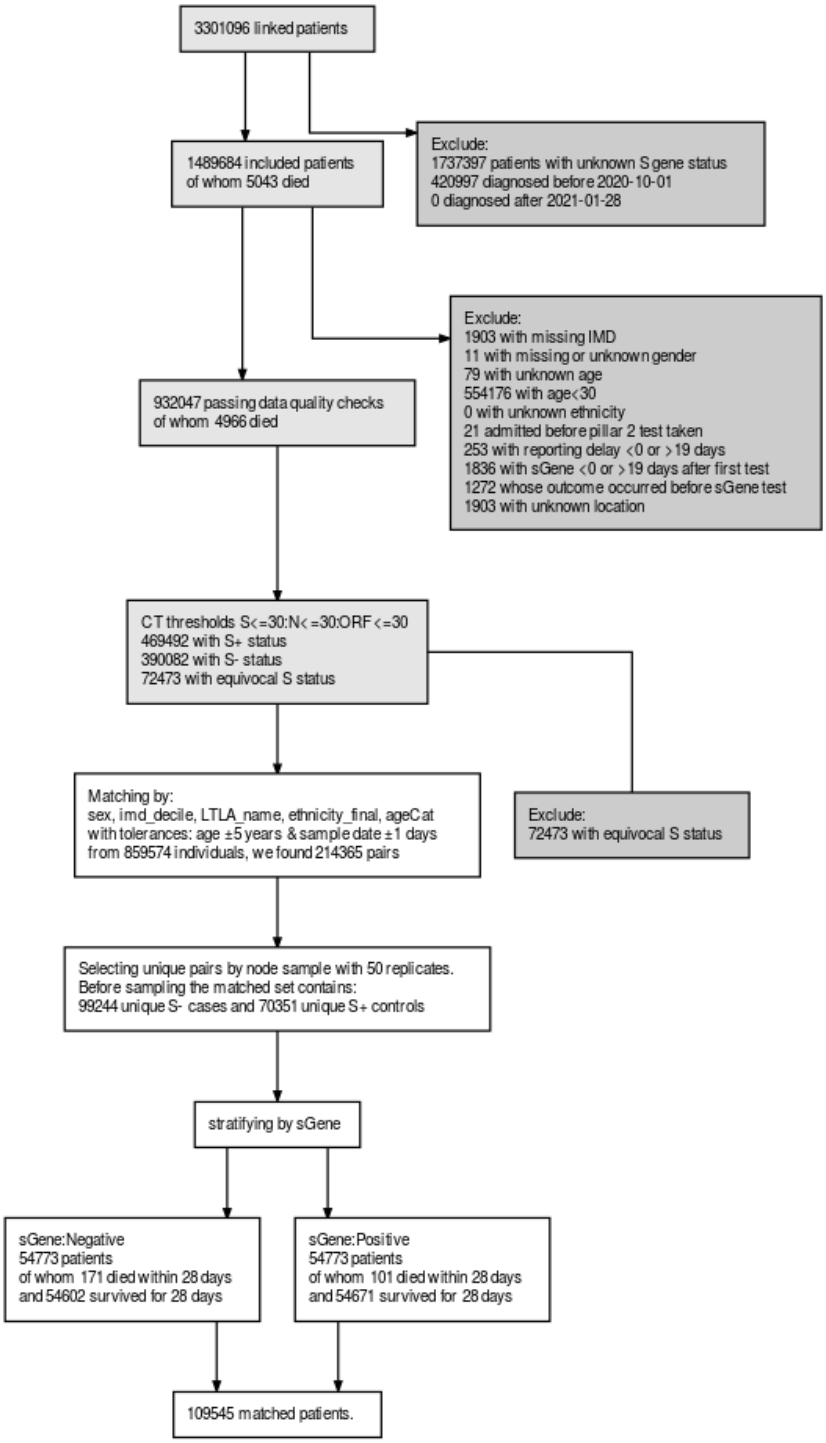

Figure 1 - the sample selection algorithm. In the matching process we randomly sample to create 50 relicates. In this figure we have given average figures for the numbers of patients in each arm of the study. LTLA is geographical location as lower tier local authority, IMD is index of multiple deprivation.

Table 1 - S-gene positive (control) & S-gene negative (case) matched case controls based on age, ethnicity, gender, index of multiple deprivation, geography and specimen date (not shown). For comparison the subset of cases and controls who died are presented in the right 3 columns. IMD is index of multiple deprivation.

category	value	S pos N	S pos %age	S pos mean (SD)	S neg N	S neg %age	S neg mean (SD)	Died N	Died %age	Died mean (SD)
Age		54773		46.3 (±11.0)	54773		46.3 (±10.9)	272		66.4 (±13.9)
Age by category	30-59	48372	88.3%		48372	88.3%		85	31.4%	
	60-69	4963	9.1%		4963	9.1%		76	28.0%	
	70-79	1173	2.1%		1173	2.1%		63	23.2%	
	80+	265	0.5%		265	0.5%		47	17.3%	
Ethnicity	White	45568	83.2%		45568	83.2%		237	87.1%	
	Other	1172	2.1%		1172	2.1%		1	0.4%	
	Afro-carribbean	993	1.8%		993	1.8%		2	0.7%	
	Asian	6914	12.6%		6914	12.6%		32	11.8%	
	Unknown	127	0.2%		127	0.2%				
Gender	Female	29297	53.5%		29297	53.5%		99	36.5%	
	Male	25477	46.5%		25477	46.5%		172	63.5%	
IMD	1	4992	9.1%		4992	9.1%		21	7.7%	
	2	9396	17.2%		9396	17.2%		68	25.0%	
	3	7232	13.2%		7232	13.2%		42	15.4%	
	4	6237	11.4%		6237	11.4%		24	8.8%	
	5	5335	9.7%		5335	9.7%		28	10.3%	
	6	4381	8.0%		4381	8.0%		27	9.9%	
	7	4409	8.0%		4409	8.0%		15	5.5%	
	8	4322	7.9%		4322	7.9%		17	6.2%	
	9	4350	7.9%		4350	7.9%		16	5.9%	
	10	4119	7.5%		4119	7.5%		14	5.1%	
N gene CT		54773		21.2 (±4.2)	54773		19.0 (±4.4)	272		18.2 (±4.3)
Region	East of England	3627	6.6%		3629	6.6%		12	4.4%	
	London	8877	16.2%		8877	16.2%		24	8.9%	
	Midlands	10515	19.2%		10527	19.2%		63	23.2%	
	North East and Yorkshire	10673	19.5%		10681	19.5%		59	21.8%	
	North West	14685	26.8%		14665	26.8%		86	31.7%	
	South East	5099	9.3%		5101	9.3%		20	7.4%	
	South West	1297	2.4%		1293	2.4%		7	2.6%	
S gene	Positive	54773	100.0%					101	37.1%	
	Neg Post B.1.1.7				54773	100.0%		171	62.9%	
Status	Dead <28 days	101	0.2%		171	0.3%		272	100.0%	
	Other	54672	99.8%		54602	99.7%				

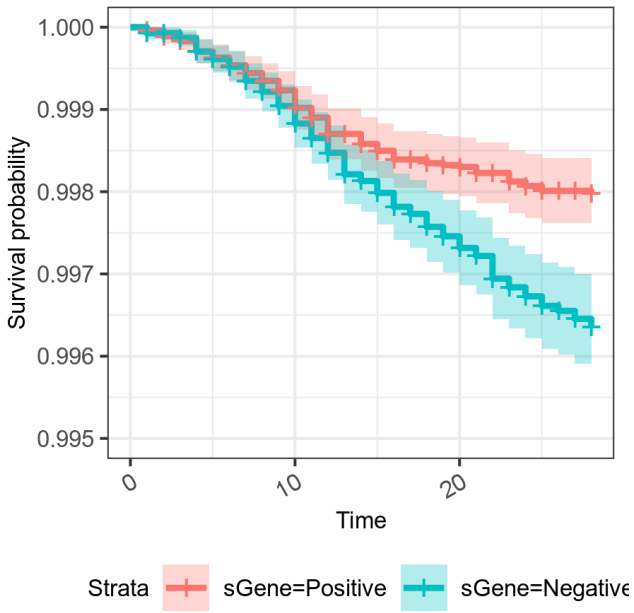

Figure 2 - A Kaplan-Meier survival curve for S-gene positive (previously circulating variants) and S-gene negative (VOC-202012/1) infections in the UK. The y axis has been clipped as mortality in both groups is low.

Table 2 - Hazard ratios for death given an S-gene negative test result versus deaths given a S-gene positive result (reference category). Hazard ratios greater than one are indicative of an increased rate of death due in infections compatible with VOC-202012/01. In the first model we look at the S-gene status as an indicator with age as a covariate, in the second model we include variability in the N gene CT value measured on original specimen as a continuous predictor, which explains some but not all of the hazard increase observed due to S-gene negativity.

Model	Predictor	Value	Hazard rate (95% CI)	p value
S gene + age	S gene status	Positive (ref)	—	—
		Negative	1.7 (1.3 – 2.2)	<0.001
	Age		1.1 (1.1 – 1.1)	<0.001
S gene + N gene CT	S gene status	Positive (ref)	—	—
		Negative	1.4 (1.1 – 1.8)	0.021
	N gene CT value		0.9 (0.9 – 0.9)	<0.001

The case matching design controls for most potential biases including variations in hospital capacity, as it pairs patients by demographics, geography and time of testing. We investigated other further potential biases that may be present. There is no evidence for asymmetric delays in time from test to admission shown in figure 3 panel A. Figure 3 panel B demonstrates that the paired cases we investigated are spread throughout time, but concentrate around the end of December 2020 and beginning of January 2021. As the ratio of S-gene negative and S-gene positive cases changes over this period, in the early stages it is hard to find S-gene negative cases for S-positive controls, and in the later stages hard to find S-positive controls for S-negative cases,

resulting in the bulk of our matched cases being during the time of transition from S-gene positive to S-gene negative variant.

We note in table 1 and in figure 3 panel C that CT values for the N-gene are lower in S-gene negative cases than in S-gene positive cases, and this effect is potentiated in those who died. Low values for N-gene cycle threshold imply the viral load in patients at the time of sampling were higher. This could be regarded either as a source of bias or as a feature of S-gene negative infection. If we interpret it as a source of bias, we can control for N-gene CT value in the Cox proportional hazards model (in table 2 - second model) which shows a reduction in the overall hazard of S-gene negativity to 1.4 (95% CI 1.1 - 1.8).

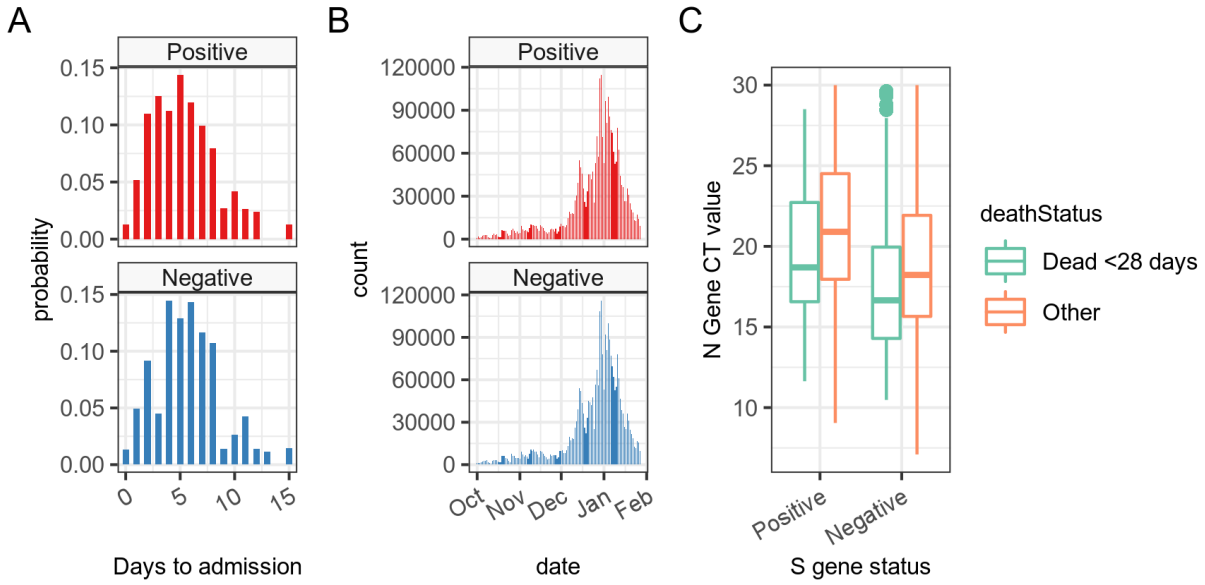

Figure 3 - Investigation of biases in the case control arms. In panel A we see no discernable pattern in the delay between specimen and admission in patients who subsequently died. In panel B we see the date distribution of the specimens in our matched pairs of S-gene positive and negative cases - concentrated in December and January. In panel C we see median CT values for the N gene for both S-gene positive and negative cases, demonstrating lower CT values in the initial tests of those who subsequently died, and in those with S-gene negative infections.

Sensitivity analysis

In preliminary work on this topic we calculated the hazard rate to be marginally higher than the estimate presented here [20,21]. This was in part due to the data available, in part down to assumptions we made during the definition of S-gene negative cases and in part due to the degree of tolerance we allowed when matching cases by age or by date of specimen.

Figure 4 shows the estimates of hazards related to altering those assumptions. In panel A we see that reducing the CT value threshold for identifying a particular gene leads to a reduction in the central estimate of hazards. Lower CT thresholds reduce the number of certain positive and negative cases and increase the number of equivocal cases, leading to an effective reduction in

overall case numbers. There appears to be a marginal, non-significant rise at a CT threshold of 30; we choose this as our central estimate since this is a standard CT threshold for most tests and is employed by most laboratories.

Panel B shows the effect of allowing age mismatched between cases and controls. This does not have the same systematic bias as sample date and the mean age difference between cases and controls is less than 0.005 years. Age is a strong predictor of mortality in COVD-19, so we might expect some potential bias, however we control for this when calculating the hazard (see Table 2).

Panel C shows the effect of allowing greater time between sample dates in matched cases and controls, increasing the temporal uncertainty and diluting the hazard. Because of the change in prevalence from S-gene positive to S-gene negative temporal uncertainty leads to a systematic bias (Panel D) with S-negative cases generally being after S-positive controls. Given that over the study period cases were exponentially increasing this risks introducing a bias if hospital capacity begins to be exceeded over time. For this reason we aimed to minimise the sample date tolerance, trading off the reduction in bias against the variance introduced by the reduced case numbers resulting from tight matching criteria.

Despite the differences between all the combinations investigated, all studies report a statistically significant rise in mortality hazard imparted by the variant, suggesting a real effect, and the vast majority of central estimates are within the range of 1.5 to 1.7.

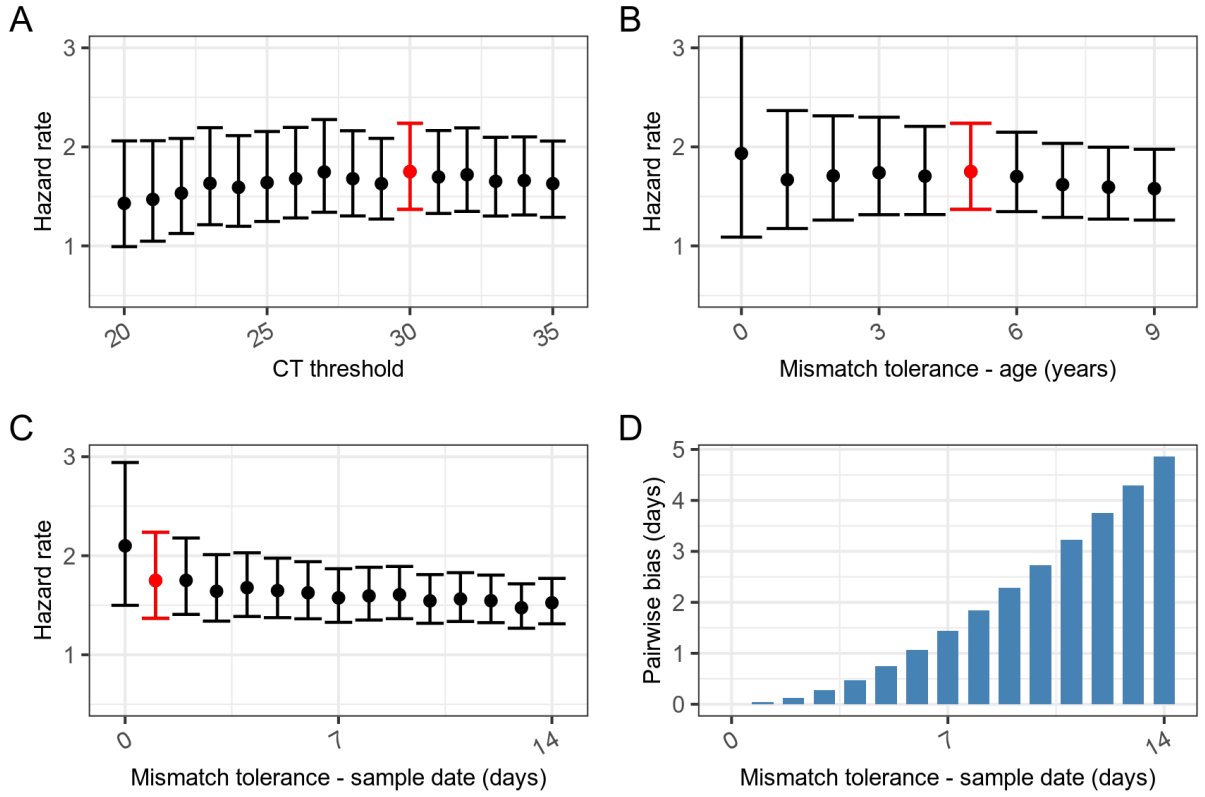

Figure 4 - Sensitivity analysis. Panel A - central HR estimates are mildly affected by the threshold for CT values used in definition of gene positivity. Panel B - allowing greater flexibility in age

when selecting cases leads to small changes in the associated HR estimate, and is included as a covariate in the Cox model. Panel C - allowing greater flexibility in sample date when selecting paired cases leads to slightly lower estimates of hazard, however because the prevalence of cases and controls changes over time, it also results in increasing systematic bias between cases and controls - panel D, where in the most permissive matching S-negative cases could be on average up to 5 days later that S-positive controls. In panels A,B & C the red bar indicates our default assumptions (CT < 30; age tolerance ±5 years; sample date tolerance ±1 day) from the rest of the paper.

Discussion

VOC-202012/1 infections (as measured by S-gene negativity) lead to an elevated risk of death ($p<0.001$) in people testing positive for COVID-19 in the community. The increased hazard varies between 1.3 and 2.2 over and above other variants, which translates to a 30%-120% increase in mortality, with the most probable HR estimate of 1.7 (70% increase).

The matched cohort approach controls for many of the biases we are aware of. In particular, mortality is affected by how many cases require intensive care in a hospital setting [13]; increasing numbers of cases in the study period (1st Oct 2020 - 21st Jan 2021), compounded by staff absenteeism due to infection with or isolation due to contact with infected individuals has placed intense strain on hospital services and a reduction in the staff to patient ratio. This may have affected mortality and is a potential source of bias. This is controlled for by matching individuals by administrative region and time of positive test (within 1 day), which constrains pairs to receive the same place and time, and we assert a similar level of care. Age related mortality is controlled for by matching by age (within 5 years), but is also controlled for by the proportional hazards model employed.

This is a community based study. We do not have information about the S-gene status of patients in hospital. Testing based in the community (Pillar 2) covers a younger age group and hence represents less severe disease than cases detected through hospital based testing (Pillar 1). In cases detected in the community, death remains a comparatively rare outcome, compared to in-hospital identified cases. Our study only includes approximately 8% of the total deaths occurring during the study period. Of all coronavirus deaths, approximately 26% occur in individuals who have had a Pillar 2 test, and only 30% of these have S-gene data [21]. Whether the increase in mortality in community cases is also observed in elderly patients, or hospitalised cases, remains to be seen.

We cannot exclude a selection bias. Pillar 2 testing is largely self selected, or driven by contact tracing. There remains a potential bias if there were a higher proportion of undetected asymptomatic cases in S-gene negative infections than in S-gene positive infections. In this event, VOC-202012/1 cases may be at a more advanced stage of infection when detected, and have a higher apparent mortality. This could be consistent with the lower N-gene CT values observed in S-gene negative cases. Our analysis, or any retrospective study based on symptomatic cases, would not be able to detect this, however early survey data suggests that S-gene negative infections are, if anything, more likely to present for testing [22]. Addressing this potential bias requires a study design capable of detecting asymptomatic infections in S-gene positives and S-gene negatives.

Some of the increased hazard could be explained by comorbidities. There is no information about comorbid conditions in the data we analysed, although this will be partly controlled by age, ethnicity and index of multiple deprivation. It is possible that people with certain comorbidities are both more susceptible to infection with VOC-202012/1 and have a higher mortality. This would tend to reduce the hazard attributable to VOC-202012/1 alone.

Other recent studies produce qualitatively similar estimates of the increased hazard. On the whole these studies use the same Pillar 2 data but have employed different study and analysis designs. They result in generally compatible point estimates of mortality hazard ratio (1.3 - 1.65), and the confidence intervals of these studies overlap with those described here [21]. As with our work, these estimates are being continuously re-evaluated as more data is acquired. The design of this study is well suited to determining whether there is an elevated risk of death, in an unbiased manner, but uses a comparatively small number of patients. Other study designs, involving the use of wider unpaired samples, may be better at quantifying the absolute increase in risk, but with more potential for bias [23].

Conclusion

The variant of concern, in addition to being more transmissible, appears to be more deadly. We expect this to be due to changes in its phenotypic properties associated with multiple genetic mutations [24], and see no reason why this finding would be specific to the UK. This concerning development, borne out in epidemiological analyses, implies that there will be an increase in the rate of serious cases requiring hospital attention. At the time of writing (31/01/2021) the national lockdown appears to be effective at reducing the transmission rate in the UK, but control of the outbreak has been made more difficult by the proliferation of the new variant. Clinicians at the front line should be aware that a higher mortality rate is likely even if quality of practice remains unchanged. This has broader implications for any vaccination allocation policy designed to reduce mortality in the late-middle age-groups typical of the community identified cases in Pillar 2.

The question remains whether excess mortality due to VOC-202012/1 will be observed in other population groups, particularly the elderly, care home residents, and those with other comorbidities who generally present directly to hospitals as emergencies. Hospital-based studies require a mechanism to distinguish emerging variants from previously circulating variants, currently this is only done through genotyping. Due to the effort involved, the proportion of genotyped samples representing hospitalised cases remains low, and we argue that PCR tests that specifically target VOC-202012/1 mutations should be more widely used.

Moreover, the emergence of VOC-202012/1 and its mutations (including E484K), combined with other variants of concern including those identified in Brazil and South Africa [25] highlights the capacity of SARS-CoV-2 to rapidly evolve new phenotypic variants, with vaccine escape mutants being a real possibility [26]. Our method helps characterise the clinical presentation and outcome of one new variant, but is generalisable to others, given sufficient amounts of informative data. However, assessment of the clinical outcomes of multiple circulating phenotypic variants requires scalable technology capable of identifying substantial numbers of cases due to emerging variants (e.g. broad PCR assay panels targeting variant foci [27]) and robust collection of outcome data.

The effect of time, location, case rates, age and treatment pathways have been controlled for in this study, but are important factors to understand if we are to improve future outcomes. Future

work on the relative impact of these may allow for better targeting of resource allocation [28], vaccine distribution strategies and, in the future, relaxation of restrictions.

References

1 Wise J. Covid-19: New coronavirus variant is identified in UK. *BMJ* 2020;**371**:m4857. doi:10.1136/bmj.m4857

2 Investigation of novel SARS-CoV-2 variant: Variant of Concern 202012/01 - Technical briefing 4.

3 Prime Minister announces Tier 4: 'Stay At Home' Alert Level in response to new COVID variant. GOV.UK. https://www.gov.uk/government/news/prime-minister-announces-tier-4-stay-at-home-alert-level-in-response-to-new-covid-variant (accessed 30 Jan 2021).

4 Risk related to the spread of new SARS-CoV-2 variants of concern in the EU/EEA – first update. Eur. Cent. Dis. Prev. Control. 2021.https://www.ecdc.europa.eu/en/publications-data/covid-19-risk-assessment-spread-new-variants-concern-eueea-first-update (accessed 30 Jan 2021).

5 CDC. US COVID-19 Cases Caused by Variants. Cent. Dis. Control Prev. 2020.https://www.cdc.gov/coronavirus/2019-ncov/transmission/variant-cases.html (accessed 30 Jan 2021).

6 Rambaut A. Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations - SARS-CoV-2 coronavirus / nCoV-2019 Genomic Epidemiology. Virological: 2020. https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563 (accessed 30 Jan 2021).

7 Starr TN, Greaney AJ, Hilton SK, *et al.* Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell* 2020;**182**:1295-1310.e20. doi:10.1016/j.cell.2020.08.012

8 Investigation of novel SARS-CoV-2 variant: Variant of Concern 202012/01 - Technical briefing 1. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/947048/Technical_Briefing_VOC_SH_NJL2_SH2.pdf (accessed 19 Jan 2021).

9 Investigation of novel SARS-CoV-2 variant: Variant of Concern 202012/01 - Technical briefing 2. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/949639/Technical_Briefing_VOC202012-2_Briefing_2_FINAL.pdf (accessed 19 Jan 2021).

10 Investigation of novel SARS-CoV-2 variant: Variant of Concern 202012/01 - Technical briefing 3.

11 Davies NG, Barnard RC, Jarvis CI, *et al.* Estimated transmissibility and severity of novel SARS-CoV-2 Variant of Concern 202012/01 in England. *medRxiv* 2020;:2020.12.24.20248822. doi:10.1101/2020.12.24.20248822

12 Plessis L du, McCrone JT, Zarebski AE, *et al.* Establishment and lineage dynamics of the SARS-CoV-2 epidemic in the UK. *Science* Published Online First: 8 January 2021. doi:10.1126/science.abf2946

13 Wilde H, Mellan T, Hawryluk I, *et al.* The association between mechanical ventilator availability and mortality risk in intensive care patients with COVID-19: A national

retrospective cohort study. *medRxiv* 2021;:2021.01.11.21249461. doi:10.1101/2021.01.11.21249461
14 Cox DR, Oakes D. *Analysis of survival data*. CRC press 1984.
15 Shinozaki T, Mansournia MA, Matsuyama Y. On hazard ratio estimators by proportional hazards models in matched-pair cohort studies. *Emerg Themes Epidemiol* 2017;**14**:6. doi:10.1186/s12982-017-0060-8
16 R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: : R Foundation for Statistical Computing 2017.
17 Therneau TM. A package for survival analysis in r. 2020. https://CRAN.R-project.org/package=survival
18 Terry M. Therneau, Patricia M. Grambsch. *Modeling survival data: Extending the Cox model*. New York: : Springer 2000.
19 Williamson EJ, Walker AJ, Bhaskaran K, *et al.* Factors associated with COVID-19-related death using OpenSAFELY. *Nature* 2020;**584**:430–6. doi:10.1038/s41586-020-2521-4
20 Iacobucci G. Covid-19: New UK variant may be linked to increased death rate, early data indicate. *BMJ* 2021;**372**:n230. doi:10.1136/bmj.n230
21 Horby P. NERVTAG note on B.1.1.7 severity. NERVTAG https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/955239/NERVTAG_paper_on_variant_of_concern__VOC__B.1.1.7.pdf (accessed 31 Jan 2021).
22 Steel K, Davies B. Coronavirus (COVID-19) Infection Survey: characteristics of people testing positive for COVID-19 in England, 27 January 2021. Office of National Statistics https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/coronaviruscovid19infectionsinthecommunityinengland/characteristicsofpeopletestingpositiveforcovid19inengland27january2021
23 Brazauskas R, Logan BR. Observational Studies: Matching or Regression? *Biol Blood Marrow Transplant* 2016;**22**:557–63. doi:10.1016/j.bbmt.2015.12.005
24 Volz E, Mishra S, Chand M, *et al.* Transmission of SARS-CoV-2 Lineage B.1.1.7 in England: Insights from linking epidemiological and genetic data. *medRxiv* 2021;:2020.12.30.20249034. doi:10.1101/2020.12.30.20249034
25 Mahase E. Covid-19: What new variants are emerging and how are they being investigated? *BMJ* 2021;**372**:n158. doi:10.1136/bmj.n158
26 Wibmer CK, Ayres F, Hermanus T, *et al.* SARS-CoV-2 501Y.V2 escapes neutralization by South African COVID-19 donor plasma. *bioRxiv* 2021;:2021.01.18.427166. doi:10.1101/2021.01.18.427166
27 Harper H, Burridge AJ, Winfield M, *et al.* Detecting SARS-CoV-2 variants with SNP genotyping. *bioRxiv* 2020;:2020.11.18.388140. doi:10.1101/2020.11.18.388140
28 Lacasa L, Challen R, Brooks-Pollock E, *et al.* A flexible method for optimising sharing of healthcare resources and demand in the context of the COVID-19 pandemic. *PLOS ONE* 2020;**15**:e0241027. doi:10.1371/journal.pone.0241027